DOCUMENT RESUME

ED 482 139 IR 058 813

AUTHOR	Goldman, Kate Haley; Bendoly, Laura
TITLE	Investigating Heuristic Evaluation: A Case Study.
PUB DATE	2003-00-00
NOTE	11p.; In: Museums and the Web 2003: Selected Papers from an International Conference (7th, Charlotte, NC, March 19-22, 2003); see IR 058 801.
AVAILABLE FROM	For full text: http://www.archimuse.com/mw2003/papers/haley/haley.html/.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Evaluation Methods; *Heuristics; *Museums; Online Systems; Program Evaluation; Use Studies; World Wide Web
IDENTIFIERS	*Usability; *Virtual Museums; Web Site Design

ABSTRACT

When museum professionals speak of evaluating a web site, they primarily mean formative evaluation, and by that they primarily mean testing the usability of the site. In the for-profit world, usability testing is a multi-million dollar industry, while non-profits often rely on far too few dollars to do too much. Hence, heuristic evaluation is one of the most popular methods of usability testing in museums. Previous research has shown that the ideal usability evaluation is a mixed methods approach, using both qualitative and quantitative, expert-focused and user-focused methods. But some within the online museum field have hypothesized that heuristic evaluation alone is sufficient to recognize most usability issues. To date there have been no studies on how reliable or valid heuristic evaluation is for museum web sites. This is critical if heuristic evaluation is to be used alone rather than in tandem with other methods. This paper focuses on work being done at the Atlanta History Center as a case study for the effectiveness of heuristic evaluation in a museum web site setting. It is a project currently in the beginning stages of development. The Center is applying a thorough mixed methods approach to evaluation, including heuristic evaluation. The results of this project will assess how complete and how useful a rigorous heuristic evaluation is alone and in conjunction with other methods in the development and implementation of an online educational resource. (Author)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 482 139

PAPERS
Museums and the Web 2003

Register
Workshops
Sessions
Speakers
Interactions
Demonstrations
Exhibits
Events
Best of the Web
Key Dates
Charlotte

A&MI
Archives & Museum Informatics
158 Lee Avenue
Toronto Ontario
M4E 2P3 Canada

ph: +1 416-691-2516
fx: +1 416-352-6025

info@archimuse.com
www.archimuse.com

Search A&MI

Join our Mailing List.
Privacy.

Updated: March 13, 2003

Investigating Heuristic Evaluation: A Case Study

Kate Haley Goldman, Institute for Learning Innovation, and Laura Bendoly, Atlanta History Center, USA

http:// www.ilinet.org

Abstract

When museum professionals speak of evaluating a web site, they primarily mean formative evaluation, and by that they primarily mean testing the usability of the site. In the for-profit world, usability testing is a multi-million dollar industry, while in non-profits we often rely on far too few dollars to do too much. Hence, heuristic evaluation is one of the most popular methods of usability testing in museums.

Previous research has shown that the ideal usability evaluation is a mixed-methods approach, using both qualitative and quantitative, expert-focused and user-focused methods. But some within the online museum field have hypothesized that heuristic evaluation alone is sufficient to recognize most usability issues. To date there has been no studies on how reliable or valid heuristic evaluation is for museum web sites. This is critical if heuristic evaluation is to be used alone rather than in tandem with other methods.

This paper will focus on work being done at the Atlanta History Center as a case study for the effectiveness of heuristic evaluation in a museum web site setting. It is a project currently in the beginning stages of development. The Center is applying a thorough mixed-methods approach to evaluation, including heuristic evaluation. The results of this project will assess how complete and how useful a rigorous heuristic evaluation is alone and in conjunction with other methods in the development and implementation of an online educational resource.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D Bearman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Keywords: Evaluation, Heuristic Evaluation, Usability

The Atlanta History Center as a Case Study

The Atlanta History Center (AHC) has begun a three-year education outreach initiative funded by the Goizueta Foundation to enhance their existing outreach program and web site and help to develop a first-rate distance learning program. The site in question focuses on the online publication of educational materials and resources developed by the Center for a target population of educators in schools, both classroom teachers and media specialists. Since this population is narrowly defined, yet of prime importance to museums, the project makes an ideal forum for testing heuristic evaluation in a museum setting. The Institute for Learning Innovation has been serving as the evaluator for the Goizueta Foundation distance learning project.

The project has three primary educational objectives:

1. Improve their existing web site and develop new content and features so that the site is more easily accessible to educators,
2. Ensure that the content better reflects Georgia's Quality Core Curriculum, and
3. Increase the number of teachers and students who use the web-based educational materials.

IR058813

BEST COPY AVAILABLE

A secondary objective, common to many museums, is to build such strong ties to the educational community that the number of school group visits to the physical site of the Atlanta History Center increases.

While creating the study design for the above-mentioned project, we became concerned about one of the most commonly used techniques in web site evaluation, known as heuristic evaluation. Heuristic evaluation is a usability engineering methodology where experts who are trained in usability but who are not the end users of the proposed technology project compare the proposed technology against established usability principals known as heuristics. The training time for this technique is relatively short- as little as a half-day workshop, and the cost is often lower than other possible usability techniques. Due to this accessibility, heuristic evaluation has been frequently used by museums.

While there are many factors to consider when selecting a research methodology, such as cost, sample size, and personnel, it is assumed that the techniques used must be fundamentally sound. Heuristic evaluation has become hotly debated within the human-computer interaction field due to concerns about the reliability and validity of the results that it produces. Some specialists claim that heuristic evaluation both overlooks usability problems that may cripple the ability of a person to use the program in question, while highlighting issues that the user never encounters. Previous research, such as that by Harm and Schweibenz (2001), has shown that the ideal usability evaluation is a mixed-methods approach, using both qualitative and quantitative, expert-focused and user-focused methods. But some within the online museum field have hypothesized that heuristic evaluation alone is sufficient to recognize most usability issues. To date there have been no studies on how reliable or valid heuristic evaluation is for museum web sites. This is critical if heuristic evaluation is to be used alone rather than in tandem with other methods.

Using the current project at the Atlanta History Center as a case study, we saw an opportunity to further investigate the issue of reliability and validity in using heuristic evaluation for museum web sites. This paper will outline our proposed techniques and current thinking; as the project develops we expect these techniques to evolve.

Why evaluate at all?

As a point of reference, it is useful to step back from the AHC project and review the goals and methodologies of both traditional museum evaluation and the developing field of museum web site evaluation. Evaluation is used to urges us to clarify our goals and accomplish our objectives. If we are able to define what we intend to do, we are more likely to achieve our goals, increase the museum's responsiveness to the community, avoid false assumptions about our visitors, and save time and money. Evaluation can be scary, because a project with unclear objectives and no evaluation can always be described as successful. This is perhaps best stated by the Flying Karamazov Brothers who said, "If you don't know where you're going, any road will get you there."

A quick review and comparison of traditional museum evaluation and museum web site evaluation is covered in Table 1. Audience research is done by some institutions on a regular cyclical basis, by some others who have done no other research and need a starting point or by those are beginning a new initiative or strategic plan. Audience research provides demographic information and other basic visitor information and is often done on the internet through log files analysis and surveys.

Traditional museum evaluation is made up of four types, not including the above mentioned audience research. Front-end evaluation typically occurs during the initial planning phase of project development and provides information about visitors' interest, expectations, and understanding of proposed topics for a program. Formative evaluation takes place while a project is in development and construction. It provides feedback on the effectiveness of a project, and its

components -- feedback which allows developers to make informed decisions as they continue to build the project. Remedial evaluation is generally conducted after a project is available to public. This type of evaluation focuses on determining changes which need to be made to the program to improve it. Summative evaluation is conducted after an exhibit or program is completed, and it seeks to determine the extent to which exhibit or program goals were met.

Museum Evaluation	Museum Web Site Evaluation	Web Site Methodologies
Audience Research	Audience Research	Log Files Analysis, Surveys
Front-End	Usability	User Testing and Think-aloud Protocols, Heuristic Evaluation, Surveys
Formative	Usability	User Testing and Think-aloud Protocols, Heuristic Evaluation, Surveys
Remedial	Usability	User Testing and Think-aloud Protocols, Heuristic Evaluation, Surveys
Summative	??????	??????

Table 1: Evaluation Types and Methods

Usability testing is a standard piece of the larger development lifecycle throughout the technology industry and has been carried over into the field of museum technology development. Usability is currently the main focus for formative, remedial and even front-end evaluation. Although usability is extremely important and is the focus of this current project, the fact that a project or program is usable does not make it de facto valuable, or even used. The logistical and methodological difficulties of assessing the value of a project when the users are geographically scattered means that summative evaluation of museum web sites being rarely undertaken.

Background on Usability Engineering Techniques

The human-computer interaction field has developed a wide range of techniques to evaluate usability of technology projects. Techniques that are expert-based are known as usability inspection techniques. For-profit companies often choose expert-based methods over user-based methods because of the high costs of doing laboratory tests with end-users.

Heuristic evaluation is one of the most informal methods of usability inspection, meaning it is based on rules of thumb and the skills of the evaluators. In heuristic evaluation, the evaluators may be non-experts who have received some training in usability principles. Since this is a less formal method which avoids using a full set of controls or specified personnel lower costs are incur than in formal testing. To quote Mack and Neilsen,

> Usability engineering activities are often difficult to justify and carry out in a timely way, but many activities can be done quickly and cheaply, and produce useful results. The methodology decision ...turn less on what is "correct" than on what can be done within development constraints. After all, with sufficient resources we would likely simply aim for rapid prototyping and end-user testing.
>
> *(Mack and Nielsen, 1994)*

Although other usability inspection techniques are rarely used in the museum

field, we will briefly describe them below in order to give a sense of what could be used or adapted as a technique for our field. The majority of these are designed for designers and developers in the formative development period of a project, rather than the front-end or remedial stage.

Possible Usability Inspection Methods:

1. Guideline Review

Project is checked to determine conformity to a list of usability guidelines. Comprehensive sets can contain more than a thousand guidelines, and require skilled expertise. They are considered a mix of heuristic evaluation and standards inspections.

2. Standards Inspections

An expert in a particular type of interface inspects the product based on guidelines for that specific product range.

3. Cognitive Walkthroughs

Exploration focused inspection focused on one feature of usability- the ease of learning. This might be a useful goal for a complex software product, but for a public web-site a more common goal is ease of use. Ease of use would mean a first-time user could navigate and accomplish his or her objective easily, as opposed to finding it easy to become an expert of a more complex system.

4. Pluralistic Walkthroughs

Group meetings with users, developers and human interaction personnel walk through user scenarios, documenting each step of the scenario and discussing implications.

5. Consistency Inspections

Inspections by designers and developers across multiple projects, ensuring that the projects have consistent design elements and usability. For instance, as multiple designers may work on separate functions of a museum web site, a consistency review would evaluate the congruity of the different sections or how well each section complies with ADA guidelines.

6. Formal Usability Inspections

Inspection method similar to software code inspections. designed to discover and report a large amount of data efficiently. Inspectors take on user roles and work through prescribed scenarios.

7. Feature Inspections

Focuses on whether the project functions as developed meet the needs of the intended end users. In traditional evaluation, this would be a part of summative evaluation.

Reliability and Validity Issues in Heuristic Evaluation

Reliability is the consistency or stability of a measure from one test to the next. Repeated measures of a static item using a reliable measure should end in identical or similar results. Validity is a term used to describe whether a measure accurately measures what it is supposed to measure. For instance, it is hotly

debated whether SAT scores accurately assess college achievement. If SATs did accurately assess achievement, they would be a valid measure.

Studies that bring the reliability of inspection methods include two studies by Rolf Molich. In the first study, he asked four commercial usability laboratories to carry out usability tests on a calendar program that was commercially available. One laboratory found as few as 4 problems, another found as many as 98. The biggest concern, however, is that only one problem was found by all four team and over 90% of the problems found by each team were found by that team alone. The second follow-up study had similar results- there was little inter-rater reliability.

The validity of usability inspection methods should be easier to address- the pertinent question asks how predictive are these methods of end-user problems? Studies on that question have been completed outside of the museum field. Karat (1994) reports on the results of several such studies. A study by Desurvire (1994) compared heuristic evaluation and an automated cognitive walkthrough to laboratory tests with end users. The system in question was not a web site, but a telephone system that completed six basic tasks. Table 2 below contrasts the results of the laboratory data with end users and the data collected using inspection methods.

Method	Evaluators	Problems That Did Occur	Potential Problems	Improve-ments
Lab	Observed with users	25	29	31
Heuristic Evaluation	Experts	44%	31%	77%
	Software developers	16%	24%	3%
	Nonexperts	8%	3%	6%
Cognitive Walkthrough	Experts	28%	31%	16%
	Software developers	16%	21%	3%
	Nonexperts	8%	7%	6%

Table 2: Prediction Rate of End-User Problems

The top line in this table indicated the number of usability problems and interface improvement ideas that were observed during user testing in the laboratory. The remaining part of the table shows the percentage of these problems and improvement ideas found by the evaluators using either heuristic evaluation or cognitive walkthrough. (Source: Desurvire 1994)

In the study above, experts were able to predict at best 44 percent of the usability problems identified by the end users. The table above does not express variance in the problems that occur. Some problems users encounter are relatively minor and others prevent the user from completing major tasks. Desurvire dealt with this issues by asking each participant to assign Problem Severity Codes to the problems uncovered. The table displaying these results is reproduced below. Note that experts were able to detect 80% of the minor problems or annoyances but only 29% of the problems that caused task failure.

		Problem Severity Code (PSC)		
Method	**Evaluators**	**Minor Annoyance/ Confusion**	**Problem Caused Error**	**Problem Caused Task**

BEST COPY AVAILABLE

				Failure
Lab	Observed with users	5	3	17
Heuristic Evaluation	Experts	80%	67%	29%
	Software developers	40%	0%	12%
	Nonexperts	20%	0%	6%
Cognitive Walkthrough	Experts	40%	67%	18%
	Software developers	0%	0%	12%
	Nonexperts	20%	0%	6%

Table 3: Prediction Rate of End-User Problems by Severity of Problem

The Top line in this table indicated the number of usability problems in three severity categories that was observed during user testing in the laboratory. The remaining part of the table shows the percentage of the problems in each of the three categories found by evaluators using either heuristic evaluation or cognitive walkthrough. (Source: Desurvire 1994)

These results raise serious questions about the validity of heuristic evaluation- about the ability of the technique to predict end-user errors. Missing any error that regularly leads to task failure is highly problematic. Worse yet, using heuristic evaluation as the sole usability technique would result in 70% of the errors that cause task failure going undetected in this example. In addition, many interface errors found by the experts using heuristic evaluation are false positives- meaning they find errors that don't actually impact the end-user, wasting development resources on what might not really be a problem.

Still, these results were gathered by a system unlike that used to evaluate museum web site. Perhaps the nature of the medium (museum web sites) allows us to use heuristic evaluation to detect a higher rate of error. Our study aims to replicate this experiment with the AHC web site.

Research Design for AHC Project

In order to test the reliability of the heuristic evaluation methodology, we will use multiple methodologies, including both heuristic evaluation as well as user testing with think-aloud protocols. These two types of methodology are quite different. Think-alouds are a user-focused methodology where we ask the user to talk-aloud while interacting with the technology, therefore hopefully revealing the conscious cognitive processes of the user. With this technique, the interplay between thought and action is revealed by the user, rather than assumed by the researcher.

Within usability engineering, an iterative design structure is critical, and the most complete designs incorporate a cyclical process of inspection methods and user testing at different point within the evaluation process. This allows a set of checks so that the solution to a interface problem does not create increased errors in other functions. For the purposes of this experiment, each technique will be performed on the exact same version of the web site. (In a typical design structure, end-user testing would occur after changes from the heuristic evaluation had already been incorporated into the web site.) For AHC project itself, there will be several iterations of evaluation that are not a part of this experiment.

In each of the methodologies used, we will develop scenarios or tasks for the

experts or end-users to complete. There are advantages and disadvantages to using the scenario approach. If carefully constructed, the scenarios can assist participants in focusing their efforts on specific interface elements. On the other hand facilitating a more open-ended inquiry will emulate the way most users experience a site- through intuitive exploration. Testers will usually then form their own scenarios with which to make sense of a site. Given that the AHC project is only one piece of a much larger site, we opt to control the scenarios. Complexity of the scenario can at times change the usability issues found, but as the interface here will be fairly straightforwardly task oriented we do not anticipate this to be a mitigating factor.

Below we will lay out the specific processes for each methodology.

Heuristic evaluation

The first step in heuristic evaluation is to decide which set of heuristic principles to use. There are many different types of usability principles. Some of the standard ones were developed by Neilsen and others in the early 1990s. (See Tables 4 & 5) By combining the principles from several different sets, we will develop a set of usability heuristics for the AHC project.

Simple and natural dialogue
Speak the users' language
Minimize the users' memory load
Consistency
Feedback
Clearly marked exits
Shortcuts
Precise and constructive error messages
Prevent errors
Help and documentation

Table 4: Example of Usability Principles by Molich and Neilsen (1990)

Visibility of system status
Match between system and the real world
Use control and freedom
Consistency and standards
Error prevention
Recognition rather than recall
Flexibility and efficiency of use
Aesthetic and minimalist design
Help users recognize, diagnose and recover from errors
Help and documentation

Table 5: Example of Usability principles by Neilsen (1994)

For the actual process we will recruit 6 evaluators. Some studies show a benefit to evaluators working in teams, while other studies show a concern that teams "filter out" valid issues. To reap the most benefit, two evaluators will work together while the rest will work individually. Evaluators will be museum professionals who are unrelated to the project at the Atlanta History Center. In order to test the "quick-to-learn" claim of heuristic evaluation, we will not be usability experts. (There is no certification for the usability profession at this time. Within the field, the expert status normally is seen as obtained after 7 years in the field.)

Since the evaluators will not be usability experts, but museum professionals, training will first be given on heuristic evaluation, including both the process and the specific principles for this evaluation. Evaluators will not be familiar with the system itself and may or may not be familiar with the proposed types of users (generally classroom teachers, but also possibly media coordinators and students), types of tasks that system users will be trying, and the contexts involved. Training will be provided to try to set the evaluator into the users' shoes. Evaluators will then be ask to imagine several scenarios while using the site. All scenarios will be described without screen-shots or specificities that would bias the evaluator in how they might approach the site. Evaluators will have an hour or more to complete the evaluation, and will be asked to resist discussing their results with others while moving through the scenarios. We will suggest that evaluators complete each scenario twice, once to gather a rough idea of the problems, and then revisit the scenario to link those problems specifically to the defined heuristic principles. Evaluators will be asked to describe in writing each of the specific issues that arise.

After the formal evaluation, a debriefing session will be held to discuss the characteristics of the site, and identify any possible alternate approaches if critical issues arise. After the brainstorming session, evaluators will be asked to rate the severity of the problems they encounter. Severity rating assists developers to prioritize the changes needed in a project.

Neilsen's severity rating is made up of three factors:

1. The frequency with which the problem occurs: Is it common or rare?
2. The impact of the problem if it occurs: Will it be easy or difficult for the users to overcome?
3. The persistence of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

Neilsen also mentions a fourth factor which he does not directly add to the others- one of market impact. He points out that certain types of usability problems can have a 'devastating effect" on the usage of a project, even if the problem is supposedly easy to overcome.

We will use an alternative system by Desurvire (1994) for severity ratings, which splits the ratings phase into two different three point scales. The first scale, the Problem Severity Code (PSC) rates the error severity as follows:

1. Minor annoyance or confusion
2. Problem caused error
3. Problem caused task failure

The second scale measures the attitude of the user towards the system, an extremely important variable in the likelihood of a user to continue with a system once errors have occurred. The ratings for this scale are below:

1. Content with the system
2. Frustrated with the system
3. Exasperated with the system

At times it is difficult to get useful severity estimates from evaluators during the actual session, when they are mostly focused on the finding of problems, rather than on the severity of the problem and how that particular problem impedes the overall purpose of the project. His suggestion is to ask the evaluators to revisit their list of problems after the debriefing session, despite the fact that the evaluators would generally not have access to the system in question.

After gathering the severity ratings, we would do several tests of inter-rater reliability, including calculating the average correlation between the severity rating provided by any two evaluators, using Kendall's coefficient of concordance, and we would also estimate the reliability of the combined judgements by using the Spearman-Brown formula.

End-User Testing

To contrast with the Heuristic evaluation, we will also complete a round of user testing at the same point in the formative development process of the web site. We will attempt to have a minimum of 15-20 user-testing sessions. Unlike in the heuristic evaluation phase, users will work separately under the assumption that most end-users of the AHC site will be working on their own. Sessions will take place either in the History Center classrooms or within a usability laboratory. Users will be recruited through the large teacher network that has worked previously with the Atlanta History Center.

Users will be given a series of tasks and asked to work through each of them while articulating their thoughts out loud in a stream-of consciousness fashion. As with the heuristic evaluation phase, users will interact directly with the interface. With each user will be an observer/facilitator who will record users' thoughts and actions as well as use appropriate prompts to probe for further information. Sessions will be audio taped and /or videotaped for further analysis.

Analysis

During both phases of testing, data will be collected on variables task completion, error data, time to complete task, error severity, and user's attitude (the PSC and PAS scales mentioned above) based on the observation of and discussion with the end user. We will provide analysis similar to Desurvire's, doing a comparison of heuristic evaluation and end-user testing on each variable. We will also present analysis on which heuristics are cited most often. If possible, we will present a comparison on the use of evaluators individually and in teams. Finally, we will present recommendations for the use of heuristic evaluation to inspect museum web sites and suggestions for future research in this field.

References

Bailey, B. (2001) How reliable is usability performance testing? Last updated Sept 2001. Consulted August 27, 2001. http://www.humanfactors.com/downloads/sep012.htm

Desurvire, H. (1994). Faster, Cheaper!! Are Usability Inspection Methods as Effective as Empirical Testing?. In J. Nielsen and R. Mack (Ed.) *Usability Inspection Methods*. New York: Wiley & Sons, Inc, 173-199

Di Blas, N., Pai Guermand, M., & P. Paolini (2002) Evaluating the Features of Museum Websites. In D. Bearman & J. Trant (Eds.) *Museums and the Web 2002 Proceedings*. CD ROM. Archives & Museum Informatics, 2002. http://www.archimuse.com/mw2002/papers/diblas/diblas.html

Harm, I. & W. Schweibenz (2001) Evaluating the Usability of a Museum Web Site. In D. Bearman & J. Trant (Eds.) *Museums and the Web 2001 Proceedings*. CD ROM. Archives & Museum Informatics, 2001. http://www.archimuse.com/mw2001/papers/schweibenz/schweibenz.html

Karat, C., (1994). A Comparison of User Interface Evaluation Methods. In J. Nielsen and R. Mack (Ed.) *Usability Inspection Methods*. New York: Wiley & Sons, Inc, 203-230

Mack, R. & J. Nielsen, (1994). Executive Summary. In J. Nielsen and R. Mack (Ed.) *Usability Inspection Methods*.. New York: Wiley & Sons, Inc, 1-23

Nielsen, J., (1994). Heuristic Evaluation. In J. Nielsen and R. Mack (Ed.) *Usability Inspection Methods*.. New York: Wiley & Sons, Inc, 25-61

Wharton, C., Rieman, J. Lewis, C. &P. Polson, (1994). The Cognitive Walkthrough Method: A Practioner's Guide. In J. Nielsen and R. Mack (Ed.) *Usability Inspection Methods*.. New York: Wiley & Sons, Inc, 105-139

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").